

University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 5 (bases 1 to 126803)
 Waterston, R.
 Direct Submission
 Submitted (21-DEC-1999) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Jul 17, 1999 this sequence version replaced gi:477113.

COMMENT

----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@wustl.wustl.edu
 ----- Summary Statistics

 Center project name: H_NH0022N19

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and
 sequencing collaboration between the NIGRI Chromosome 7 Mapping
 Project (Eric D. Green, Director), John D. McPherson in the
 Department of Genetics (Washington University), and the Washington
 University Genome Sequencing Center. For additional information
 about the map position of this sequence, see
<http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send
<mailto:regreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPT-11 human BAC library was made from the blood of one male
 donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Fenggen, E.,
 Tateo, M., Catanesi, J.J. and de Jong, P.J. (1998) An improved
 approach for construction of bacterial artificial chromosome
 libraries. Genomics 51:1-8. The clone may be obtained either from
 Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong
 and coworkers at the Roswell Park Cancer Institute
 (<http://bacpac.med.buffalo.edu>)
 VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is CTA-332P12, 200 bp overlap. The
 clone sequenced to the right is CTB-11H14, 200 bp overlap. Actual
 start of this clone is at base position 81540 of CTA-332P12; actual
 end is at base position 14809 of CTB-11H14.

FEATURES

source

Location/Qualifiers

1..126803

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="7"

/map="7q22"

/clone="RP11-22N19"

/clone_lib="RPT-11"

28..136

/rpt_family="Retroviral"

806..926

/rpt_family="MER1-type"

1018..1961

/rpt_family="L2"

2105..2330

/rpt_family="MIR"

2562..2852

repeat_region
 3032..3102
 /rpt_family="Alu"

repeat_region
 3625..3739
 /rpt_family="L2"

repeat_region
 3966..4242
 /rpt_family="Alu"

repeat_region
 4660..4731
 /rpt_family="Ga-rich"

repeat_region
 5436..5721
 /rpt_family="MIR"

repeat_region
 5788..6025
 /rpt_family="MIR"

repeat_region
 6341..6622
 /rpt_family="Alu"

repeat_region
 6738..6759
 /rpt_family="AT-rich"

repeat_region
 7699..7909
 /rpt_family="MIR"

repeat_region
 8132..8256
 /rpt_family="L1"

repeat_region
 9057..10448
 /rpt_family="MER2-type"

repeat_region
 10454..10584
 /rpt_family="MER2-type"

repeat_region
 11748..11829
 /rpt_family="MER16"

repeat_region
 11833..12128
 /rpt_family="Alu"

repeat_region
 12401..12763
 /rpt_family="Retroviral"

repeat_region
 13157..13457
 /rpt_family="Alu"

repeat_region
 13822..13920
 /rpt_family="U6"

repeat_region
 13923..14059
 /rpt_family="L1"

repeat_region
 14275..14376
 /rpt_family="MIR"

repeat_region
 16260..17099
 /rpt_family="L1"

repeat_region
 17100..17462
 /rpt_family="L1"

repeat_region
 17449..21719
 /rpt_family="L1"

repeat_region
 21760..21844
 /rpt_family="MIR"

repeat_region
 22057..22460
 /rpt_family="MALR"

repeat_region
 22566..22899
 /rpt_family="L2"

repeat_region
 23023..23097
 /rpt_family="L2"

repeat_region
 23514..23578
 /rpt_family="MIR"

repeat_region
 23596..23883
 /rpt_family="MER2-type"

STS
 2418..24342
 /db_xref="GI:454569"

repeat_region
 25915..26030
 /rpt_family="MIR"

repeat_region
 26172..26631
 /rpt_family="MALR"

repeat_region
 27175..27456
 /rpt_family="Alu"

repeat_region
 27457..27504
 /rpt_family="A-rich"

repeat_region
 27506..27782
 /rpt_family="Alu"

repeat_region
 28153..28267
 /rpt_family="MIR"

repeat_region
 28272..28311
 /rpt_family="(CTA)n"

```
repeat_region 29763. .30062
                /rpt_family="Alu"
repeat_region 30315. .30382
                /rpt_family="MaLR"
repeat_region 30494. .30563
                /rpt_family="MaLR"
repeat_region 30717. .30804
                /rpt_family="MER1_type"
repeat_region 31224. .31765
                /rpt_family="Other"
repeat_region 32610. .32885
                /rpt_family="Alu"
repeat_region 32995. .33038
                /rpt_family="T-rich"
repeat_region 33046. .34923
                /rpt_family="L1"
repeat_region 34924. .35226
                /rpt_family="Alu"
repeat_region 35227. .36291
                /rpt_family="L1"
repeat_region 36472. .36512
                /rpt_family="(TATG)n"
repeat_region 36514. .36800
                /rpt_family="Alu"
repeat_region 36977. .37133
```

```
Query Match      92.4%; Score 451.6; DB 9; Length 126803;
Best Local Similarity 98.9%; Pred. No. 3.7e-94;
Matches 454; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
Qy      21 TGTACTATTATTGGCATATAGTTTCAAAAATTCACAGAAGGGAGCCAGGTGCTCTCATTG 80
          |||
Db      69732 TGTACTATTATTGGTATATAGTTTCAAAAATTCACAGAAGGGAGCCAGGTGCTCTCATTG 69673

Qy      81 CCTAACAAAATAATGGAAATATGTATTCAATCTTAACATCTTGACATACAGTATAAAGGGC 140
          |||
Db      69672 CCTAACAAAATAATGGAAATATGTATTCAATCTTAACATCTTGACATACAGTATAAAGGGC 69613

Qy      141 ACTCAGCAAGTGCTTTTAGTTAGACTGATTTTAAATGAGTAGATTGAGGAATACCGAA 200
          |||
Db      69612 ACTCAGCAAGTGCTTTTAGTTAGACTGATTTTAAATGAGTAGATTGAGGAATACCGAA 69553

Qy      201 AACGCACACACACAGATAATTGGTTGAAAATACGAGGTTAGAACTTAACCATGAAG 260
          |||
Db      69552 AACGCACACACACAGATAATTGGTTGAAAATACGAGGTTAGAACTTAACCATGAAG 69493

Qy      261 AGGAGTAAGGCTACTTAAAGCGTTAAAACTAATTGGGTAAGGTATGGTTGACCCAGCTA 320
          |||
Db      69492 AGGAGTAAGGCTACTTAAAGCGTTAAAACTAATTGGGTAAGGTATGGTTGACCCAGCTA 69433

Qy      321 CTTCAATTTGCCCTACGATGTATATTCATTAACTAAGTCACTCACCTTCTCTGTGATGTT 380
          |||
Db      69432 CTTCAATTTGCCCTACGATGTATATTCATTAACTAAGTCACTCACCTTCTCTGTGATGTT 69373

Qy      381 GATGGTTTGGTATAGTAAACATGAGATATGATTAAAGGTGATTGAGGATAGATCAAGTG 440
          |||
Db      69372 GATGGTTTGGTATAGTAAACATGAGATATGATTAAAGGTGATTGAGGATAGATCAAGTG 69313

Qy      441 TCTGCCTAAGTAAATCTGGGNTTCAATTTTTTTTCTAG 479
          |||
Db      69312 TCTGCCTAAGTAAATCTGTGTTTCATTTTTTTTCTAG 69274
```

RESULT 3
 AC007032/c
 LOCUS AC007032 126803 bp DNA linear PRI 21-DEC-1999
 DEFINITION Homo sapiens BAC clone RP11-22N19 from 7q22, complete sequence.
 ACCESSION AC007032
 VERSION AC007032.2 GI:5523832
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 126803)
 AUTHORS Sulston, J.E. and Waterston, R.
 TITLE Toward a complete human genome sequence
 JOURNAL Genome Res. 8 (11): 1097-1108 (1998)
 MEDLINE 99063792
 PUBMED 9847074
 REFERENCE 2 (bases 1 to 126803)
 AUTHORS Maas, J., Wohldmann, P., Harper, M. and Phillips, A.
 TITLE The sequence of Homo sapiens BAC clone RP11-22N19
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 126803)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (06-MAR-1999) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 REFERENCE 4 (bases 1 to 126803)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (17-JUL-1999) Genome Sequencing Center, Washington